

OIPE

## RAW SEQUENCE LISTING

DATE: 01/15/2002

PATENT APPLICATION: US/10/029,359

TIME: 07:45:59

Input Set : A:\10806156.txt

Output Set: N:\CRF3\01152002\J029359.raw

Does Not Comply  
Corrected Diskette Needed

f2

3 <110> APPLICANT: Pharmacia AB  
 5 <120> TITLE OF INVENTION: Protein Cluster II  
 7 <130> FILE REFERENCE: 00404  
 9 <140> CURRENT APPLICATION NUMBER: US/10/029,359  
 9 <141> CURRENT FILING DATE: 2001-12-21  
 9 <160> NUMBER OF SEQ ID NOS: 6  
 11 <170> SOFTWARE: PatentIn version 3.0  
 13 <210> SEQ ID NO: 1  
 14 <211> LENGTH: 505  
 15 <212> TYPE: DNA  
 16 <213> ORGANISM: human  
 18 <220> FEATURE:  
 19 <221> NAME/KEY: CDS  
 20 <222> LOCATION: (21)..(497)  
 22 <220> FEATURE:  
 23 <221> NAME/KEY: misc\_feature  
 24 <222> LOCATION: (297)..()  
 25 <223> OTHER INFORMATION: n=a, g, c or t  
 27 <400> SEQUENCE: 1  
 28 gcaagagggga gccacggccg atg aca gaa aat tca ctt tcc gag atg gcc tct 53  
 29 Met Thr Glu Asn Ser Leu Ser Glu Met Ala Ser  
 30 1 5 10  
 32 aaa tcc tgg ctg aat ttt tta acc ttc ctc tat gga tcg gca ata ggg 101  
 33 Lys Ser Trp Leu Asn Phe Leu Thr Phe Leu Tyr Gly Ser Ala Ile Gly  
 34 15 20 25  
 36 ttt att tta ttt tct cag cta ctt agt att ttg ttg gga gaa gag ggt 149  
 37 Phe Ile Leu Phe Ser Gln Leu Leu Ser Ile Leu Leu Gly Glu Glu Gly  
 38 30 35 40  
 40 gac acc cag act aat gtt ctt cat aat gat cct cat gcg agg cat tca 197  
 41 Asp Thr Gln Thr Asn Val Leu His Asn Asp Pro His Ala Arg His Ser  
 42 45 50 55  
 44 gat gat aat gga cag aat cat cta gga gga caa atg aac ttc aat gca 245  
 45 Asp Asp Asn Gly Gln Asn His Leu Gly Gly Gln Met Asn Phe Asn Ala  
 46 60 65 70 75  
 48 gat tct agc caa cgt aaa gat gag aac aca gaa atc gct gaa aac ctc 293  
 49 Asp Ser Ser Gln Arg Lys Asp Glu Asn Thr Glu Ile Ala Glu Asn Leu  
 50 80 85 90  
 52 tat nag caa gtt aaa att ctt tgc tgg gtt atg aca ggc tct caa aac 341  
 53 Tyr Xaa Gln Val Lys Ile Leu Cys Trp Val Met Thr Gly Ser Gln Asn  
 54 95 100 105  
 56 cta cag aaa aag gcc aaa cat gtc aaa gct aca tgg gcc cag cgt tgt 389  
 57 Leu Gln Lys Lys Ala Lys His Val Lys Ala Thr Trp Ala Gln Arg Cys  
 58 110 115 120  
 60 cta aaa gta ttt ttt atg agt tca gaa gaa aat aaa gac ttc cgt gct 437  
 61 Leu Lys Val Phe Phe Met Ser Ser Glu Glu Asn Lys Asp Phe Arg Ala  
 62 125 130 135  
 64 gtg gga ttg aaa acc aaa gca ggc aga gat gag cta tac tgg aaa aca 485

## RAW SEQUENCE LISTING

DATE: 01/15/2002

PATENT APPLICATION: US/10/029,359

TIME: 07:45:59

Input Set : A:\10806156.txt

Output Set: N:\CRF3\01152002\J029359.raw

65 Val Gly Leu Lys Thr Lys Ala Gly Arg Asp Glu Leu Tyr Trp Lys Thr  
 66 140 145 150 155

505

68 att aac ctt ttc agtatggt

69 Ile Asn Leu Phe

73 &lt;210&gt; SEQ ID NO: 2

74 &lt;211&gt; LENGTH: 159

75 &lt;212&gt; TYPE: PRT

76 &lt;213&gt; ORGANISM: human

78 &lt;220&gt; FEATURE:

79 &lt;221&gt; NAME/KEY: misc\_feature

80 &lt;222&gt; LOCATION: (297)..()

81 &lt;223&gt; OTHER INFORMATION: n=a, g, c or t

83 &lt;400&gt; SEQUENCE: 2

85 Met Thr Glu Asn Ser Leu Ser Glu Met Ala Ser Lys Ser Trp Leu Asn

86 1 5 10 15

89 Phe Leu Thr Phe Leu Tyr Gly Ser Ala Ile Gly Phe Ile Leu Phe Ser

90 20 25 30

93 Gln Leu Leu Ser Ile Leu Leu Gly Glu Glu Gly Asp Thr Gln Thr Asn

94 35 40 45

97 Val Leu His Asn Asp Pro His Ala Arg His Ser Asp Asp Asn Gly Gln

98 50 55 60

101 Asn His Leu Gly Gly Gln Met Asn Phe Asn Ala Asp Ser Ser Gln Arg

102 65 70 75 80

W--&gt; 105 Lys Asp Glu Asn Thr Glu Ile Ala Glu Asn Leu Tyr (Xaa) Gln Val Lys

106 85 90 95

109 Ile Leu Cys Trp Val Met Thr Gly Ser Gln Asn Leu Gln Lys Lys Ala

110 100 105 110

113 Lys His Val Lys Ala Thr Trp Ala Gln Arg Cys Leu Lys Val Phe Phe

114 115 120 125

117 Met Ser Ser Glu Glu Asn Lys Asp Phe Arg Ala Val Gly Leu Lys Thr

118 130 135 140

121 Lys Ala Gly Arg Asp Glu Leu Tyr Trp Lys Thr Ile Asn Leu Phe

122 145 150 155

125 &lt;210&gt; SEQ ID NO: 3

126 &lt;211&gt; LENGTH: 489

127 &lt;212&gt; TYPE: DNA

128 &lt;213&gt; ORGANISM: human

130 &lt;220&gt; FEATURE:

131 &lt;221&gt; NAME/KEY: CDS

132 &lt;222&gt; LOCATION: (55)..(489)

134 &lt;400&gt; SEQUENCE: 3

136 catctaaaaa gactgatgaa gttgattgca aatgctagtc atcataaata ccag aac

137 Asn

138 1

140 aca ggt gtc act gac aaa ctc tat caa aag atg aaa att ctt tgc tgg 105

141 Thr Gly Val Thr Asp Lys Leu Tyr Gln Lys Met Lys Ile Leu Cys Trp

142 5 10 15

144 att atg aca gga cct caa aat cta gaa aaa aag atc aga cgc atc aga 153

145 Ile Met Thr Gly Pro Gln Asn Leu Glu Lys Lys Ile Arg Arg Ile Arg

*This is an amino acid sequence*  
*Xaa is at*  
*location 93.*

## RAW SEQUENCE LISTING

DATE: 01/15/2002

PATENT APPLICATION: US/10/029,359

TIME: 07:45:59

Input Set : A:\10806156.txt

Output Set: N:\CRF3\01152002\J029359.raw

```

146          20          25          30
148 gat aca tgg gcc cag ggt tgc aat aaa gcg ttg ttt atg agc tca aaa      201
149 Asp Thr Trp Ala Gln Gly Cys Asn Lys Ala Leu Phe Met Ser Ser Lys
150      35          40          45
152 gaa aat aaa gac ttc tct act gtg gga tta cac acc aaa gaa gac aga      249
153 Glu Asn Lys Asp Phe Ser Thr Val Gly Leu His Thr Lys Glu Asp Arg
154 50          55          60          65
156 aac caa ctg tcc tgg aaa ata gtt aaa gct ttt cta tat gct cat gac      297
157 Asn Gln Leu Ser Trp Lys Ile Val Lys Ala Phe Leu Tyr Ala His Asp
158          70          75          80
160 cat tat ctg gaa tac atg gat tgg ttc atg aaa gca gat gat gat ata      345
161 His Tyr Leu Glu Tyr Met Asp Trp Phe Met Lys Ala Asp Asp Asp Ile
162          85          90          95
164 tgt ata tat atc aca ttg gac aac ttg aaa tgg ctt ctc aca aac tat      393
165 Cys Ile Tyr Ile Thr Leu Asp Asn Leu Lys Trp Leu Leu Thr Asn Tyr
166      100          105          110
168 aac cct gat gaa tcc act tac ttt ggg aaa aga ttt aag cac tgc aga      441
169 Asn Pro Asp Glu Ser Thr Tyr Phe Gly Lys Arg Phe Lys His Cys Arg
170      115          120          125
172 aaa cag gac tac atg act gga gga gca gga tat gta ctg agc aaa gaa      489
173 Lys Gln Asp Tyr Met Thr Gly Gly Ala Gly Tyr Val Leu Ser Lys Glu
174 130          135          140          145
177 <210> SEQ ID NO: 4
178 <211> LENGTH: 145
179 <212> TYPE: PRT
180 <213> ORGANISM: human
182 <400> SEQUENCE: 4
184 Asn Thr Gly Val Thr Asp Lys Leu Tyr Gln Lys Met Lys Ile Leu Cys
185 1          5          10          15
188 Trp Ile Met Thr Gly Pro Gln Asn Leu Glu Lys Lys Ile Arg Arg Ile
189          20          25          30
192 Arg Asp Thr Trp Ala Gln Gly Cys Asn Lys Ala Leu Phe Met Ser Ser
193      35          40          45
196 Lys Glu Asn Lys Asp Phe Ser Thr Val Gly Leu His Thr Lys Glu Asp
197      50          55          60
200 Arg Asn Gln Leu Ser Trp Lys Ile Val Lys Ala Phe Leu Tyr Ala His
201 65          70          75          80
204 Asp His Tyr Leu Glu Tyr Met Asp Trp Phe Met Lys Ala Asp Asp Asp
205          85          90          95
208 Ile Cys Ile Tyr Ile Thr Leu Asp Asn Leu Lys Trp Leu Leu Thr Asn
209          100          105          110
212 Tyr Asn Pro Asp Glu Ser Thr Tyr Phe Gly Lys Arg Phe Lys His Cys
213      115          120          125
216 Arg Lys Gln Asp Tyr Met Thr Gly Gly Ala Gly Tyr Val Leu Ser Lys
217      130          135          140
220 Glu
221 145
224 <210> SEQ ID NO: 5
225 <211> LENGTH: 1560

```

## RAW SEQUENCE LISTING

DATE: 01/15/2002

PATENT APPLICATION: US/10/029,359

TIME: 07:45:59

Input Set : A:\10806156.txt

Output Set: N:\CRF3\01152002\J029359.raw

```

226 <212> TYPE: DNA
227 <213> ORGANISM: human
229 <220> FEATURE:
230 <221> NAME/KEY: CDS
231 <222> LOCATION: (2)..(934)
233 <400> SEQUENCE: 5
234 a gat aat gga cag aat cat cta gaa gga caa atg aac ttc aat gca gat      49
235   Asp Asn Gly Gln Asn His Leu Glu Gly Gln Met Asn Phe Asn Ala Asp
236   1          5          10          15
238 tct agc caa cat aaa gat gag aac aca gac att gct gaa aac ctc tat      97
239 Ser Ser Gln His Lys Asp Glu Asn Thr Asp Ile Ala Glu Asn Leu Tyr
240          20          25          30
242 cag aaa gtt aga att ctt tgc tgg gtt atg acc ggc cct caa aac cta      145
243 Gln Lys Val Arg Ile Leu Cys Trp Val Met Thr Gly Pro Gln Asn Leu
244          35          40          45
246 gag aaa aag gcc aaa cac gtc aaa gct act tgg gcc cag cgt tgt aac      193
247 Glu Lys Lys Ala Lys His Val Lys Ala Thr Trp Ala Gln Arg Cys Asn
248          50          55          60
250 aaa gtg ttg ttt atg agt tca gaa gaa aat aaa gac ttc cct gct gtg      241
251 Lys Val Leu Phe Met Ser Ser Glu Glu Asn Lys Asp Phe Pro Ala Val
252 65          70          75          80
254 gga ctg aaa acc aaa gaa ggc aga gat caa cta tac tgg aaa aca att      289
255 Gly Leu Lys Thr Lys Glu Gly Arg Asp Gln Leu Tyr Trp Lys Thr Ile
256          85          90          95
258 aaa gct ttt cag tat gtt cat gaa cat tat tta caa gat gct gat tgg      337
259 Lys Ala Phe Gln Tyr Val His Glu His Tyr Leu Gln Asp Ala Asp Trp
260          100          105          110
262 ttt ttg aaa gca gat gat gac acg tat gtc ata cta gac aat ttg agg      385
263 Phe Leu Lys Ala Asp Asp Asp Thr Tyr Val Ile Leu Asp Asn Leu Arg
264          115          120          125
266 tgg ctt ctt tca aaa tac gac cct gaa gaa ccc att tac ttt ggg aga      433
267 Trp Leu Leu Ser Lys Tyr Asp Pro Glu Glu Pro Ile Tyr Phe Gly Arg
268          130          135          140
270 aga ttt aag cct tat gta aag cag ggc tac atg agt gga gga gca gga      481
271 Arg Phe Lys Pro Tyr Val Lys Gln Gly Tyr Met Ser Gly Gly Ala Gly
272 145          150          155          160
274 tat gta cta agc aaa gaa gcc ttg aaa aga ttt gtt gat gca ttt aaa      529
275 Tyr Val Leu Ser Lys Glu Ala Leu Lys Arg Phe Val Asp Ala Phe Lys
276          165          170          175
278 aca gac aag tgt aca cat agt tcc tcc att gaa gac tta gca ctg ggg      577
279 Thr Asp Lys Cys Thr His Ser Ser Ser Ile Glu Asp Leu Ala Leu Gly
280          180          185          190
282 aga tgc atg gaa att atg aat gta gaa gca gga gat tcc aga gat acc      625
283 Arg Cys Met Glu Ile Met Asn Val Glu Ala Gly Asp Ser Arg Asp Thr
284          195          200          205
286 att gga aaa gaa act ttt cat ccc ttt gtg cca gaa cac cat tta att      673
287 Ile Gly Lys Glu Thr Phe His Pro Phe Val Pro Glu His His Leu Ile
288          210          215          220
290 aaa ggt tat cta cct aga acg ttt tgg tac tgg aat tac aac tat tat      721

```

## RAW SEQUENCE LISTING

DATE: 01/15/2002

PATENT APPLICATION: US/10/029,359

TIME: 07:45:59

Input Set : A:\10806156.txt

Output Set: N:\CRF3\01152002\J029359.raw

```

291 Lys Gly Tyr Leu Pro Arg Thr Phe Trp Tyr Trp Asn Tyr Asn Tyr Tyr
292 225                230                235                240
294 cct cct gta gag ggt cct ggt tgc tgc tct gat ctt gca gtt tct ttt      769
295 Pro Pro Val Glu Gly Pro Gly Cys Cys Ser Asp Leu Ala Val Ser Phe
296                245                250                255
298 cac tat gtt gat tct aca acc atg tat gag tta gaa tac ctc gtt tat      817
299 His Tyr Val Asp Ser Thr Thr Met Tyr Glu Leu Glu Tyr Leu Val Tyr
300                260                265                270
302 cat ctt cgt cca tat ggt tat tta tac aga tat caa cct acc tta cct      865
303 His Leu Arg Pro Tyr Gly Tyr Leu Tyr Arg Tyr Gln Pro Thr Leu Pro
304                275                280                285
306 gaa cgt ata cta aag gaa att agt caa gca aac aaa aat gaa gat aca      913
307 Glu Arg Ile Leu Lys Glu Ile Ser Gln Ala Asn Lys Asn Glu Asp Thr
308                290                295                300
310 aaa gtg aag tta gga aat cct tgaaagaaaa tcatgaatga acaaaggtaa      964
311 Lys Val Lys Leu Gly Asn Pro
312 305                310
314 tatgtctagc actgcactga aaaaggactt ctgcatttct gacatagaac actggaatcc      1024
316 cagtgaggaa ttctaagtga acattcctta tagaaacctt tcacatgaat gactataaac      1084
318 tgaagcttta aatgagctgt gaagtgtggt aaaatgtggt ttgatacagt aatatataaa      1144
320 tatgtctata tatatgagga acttgtgttt tttaaatggt ggccaggtag aggaactaga      1204
322 aaagagattt tgttgctgtt tttctgacca tctgtgttat tgtcactgag aaactaaaaat      1264
324 agtaaaattt ctaaaactac actgcacccat gttagtaata aacagatctg ccttaaagaa      1324
326 aagaaaaattt tagaaagaaa tattgttgtt cagtgttgtt aatatagctc aagaattgag      1384
328 tttatatattg cagtatgcta taaatgatac cccctacca caccacaca cacagttttt      1444
330 gtctaatagaa aatgttgcctg tgattattta taattggtag tttttcttcc agaagaagct      1504
332 aaaataagac tggcacttac cctgaagtgc attaataaaa ccacacttta aaatta      1560
335 <210> SEQ ID NO: 6
336 <211> LENGTH: 311
337 <212> TYPE: PRT
338 <213> ORGANISM: human
340 <400> SEQUENCE: 6
341 Asp Asn Gly Gln Asn His Leu Glu Gly Gln Met Asn Phe Asn Ala Asp
342 1                5                10                15
345 Ser Ser Gln His Lys Asp Glu Asn Thr Asp Ile Ala Glu Asn Leu Tyr
346                20                25                30
349 Gln Lys Val Arg Ile Leu Cys Trp Val Met Thr Gly Pro Gln Asn Leu
350                35                40                45
353 Glu Lys Lys Ala Lys His Val Lys Ala Thr Trp Ala Gln Arg Cys Asn
354                50                55                60
357 Lys Val Leu Phe Met Ser Ser Glu Glu Asn Lys Asp Phe Pro Ala Val
358 65                70                75                80
361 Gly Leu Lys Thr Lys Glu Gly Arg Asp Gln Leu Tyr Trp Lys Thr Ile
362                85                90                95
365 Lys Ala Phe Gln Tyr Val His Glu His Tyr Leu Gln Asp Ala Asp Trp
366                100                105                110
369 Phe Leu Lys Ala Asp Asp Asp Thr Tyr Val Ile Leu Asp Asn Leu Arg
370                115                120                125
373 Trp Leu Leu Ser Lys Tyr Asp Pro Glu Glu Pro Ile Tyr Phe Gly Arg

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/029,359

DATE: 01/15/2002

TIME: 07:46:00

Input Set : A:\10806156.txt

Output Set: N:\CRF3\01152002\J029359.raw

L:9 M:270 C: Current Application Number differs, Replaced Current Application No

L:9 M:271 C: Current Filing Date differs, Replaced Current Filing Date

L:52 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1

L:53 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1

L:105 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2